Exploring and Discovering Systematics with Precursor Data and Null Tests

Mike Jarvis

December 15, 2015

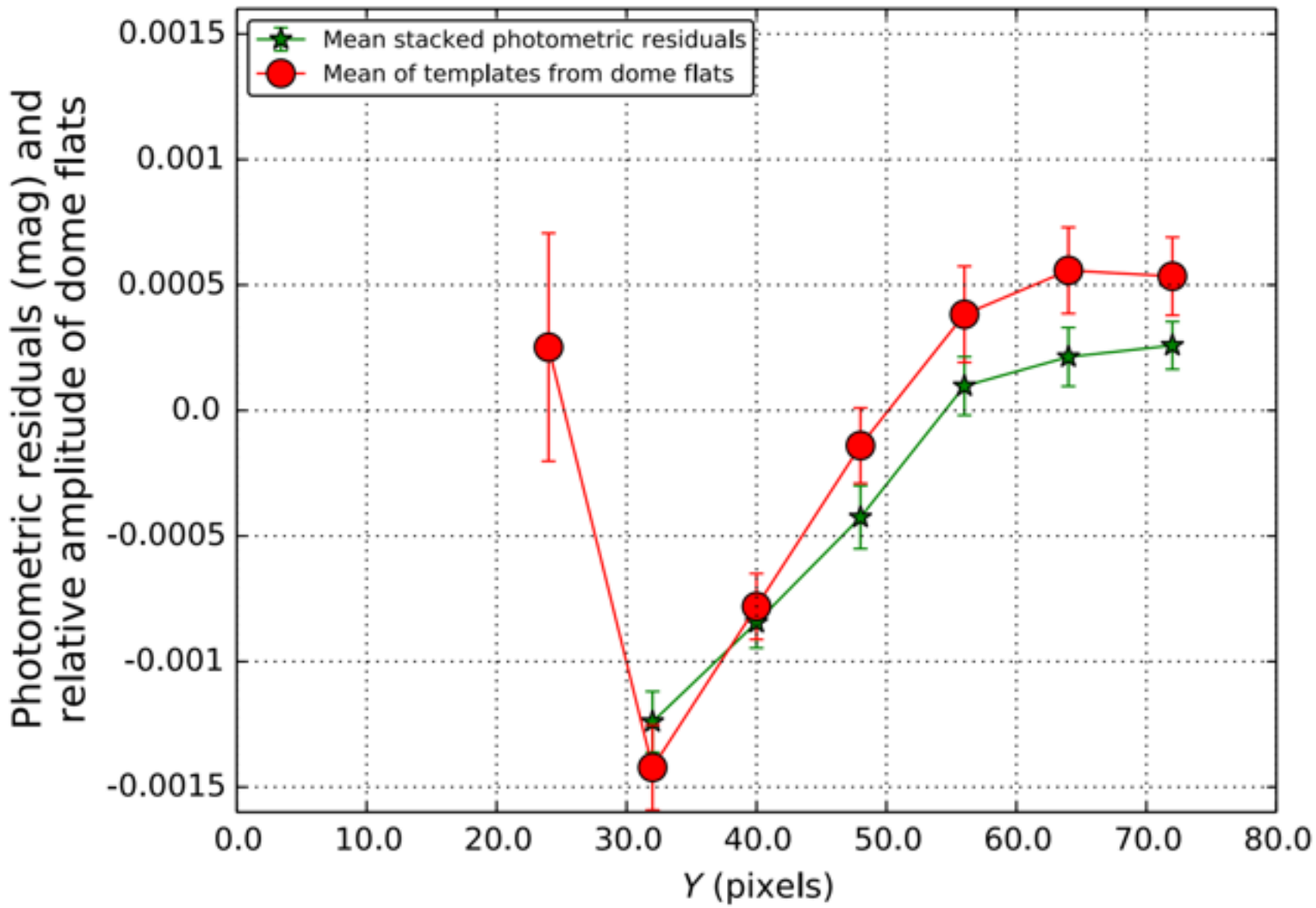
Foreground Physical Effects on
LSST Weak Lensing Science

DARK ENERGY SURVEY

What is a Null Test?

- Any statistic that should be zero in the absence of systematic errors.

Image credit: Plazas et al, 2014

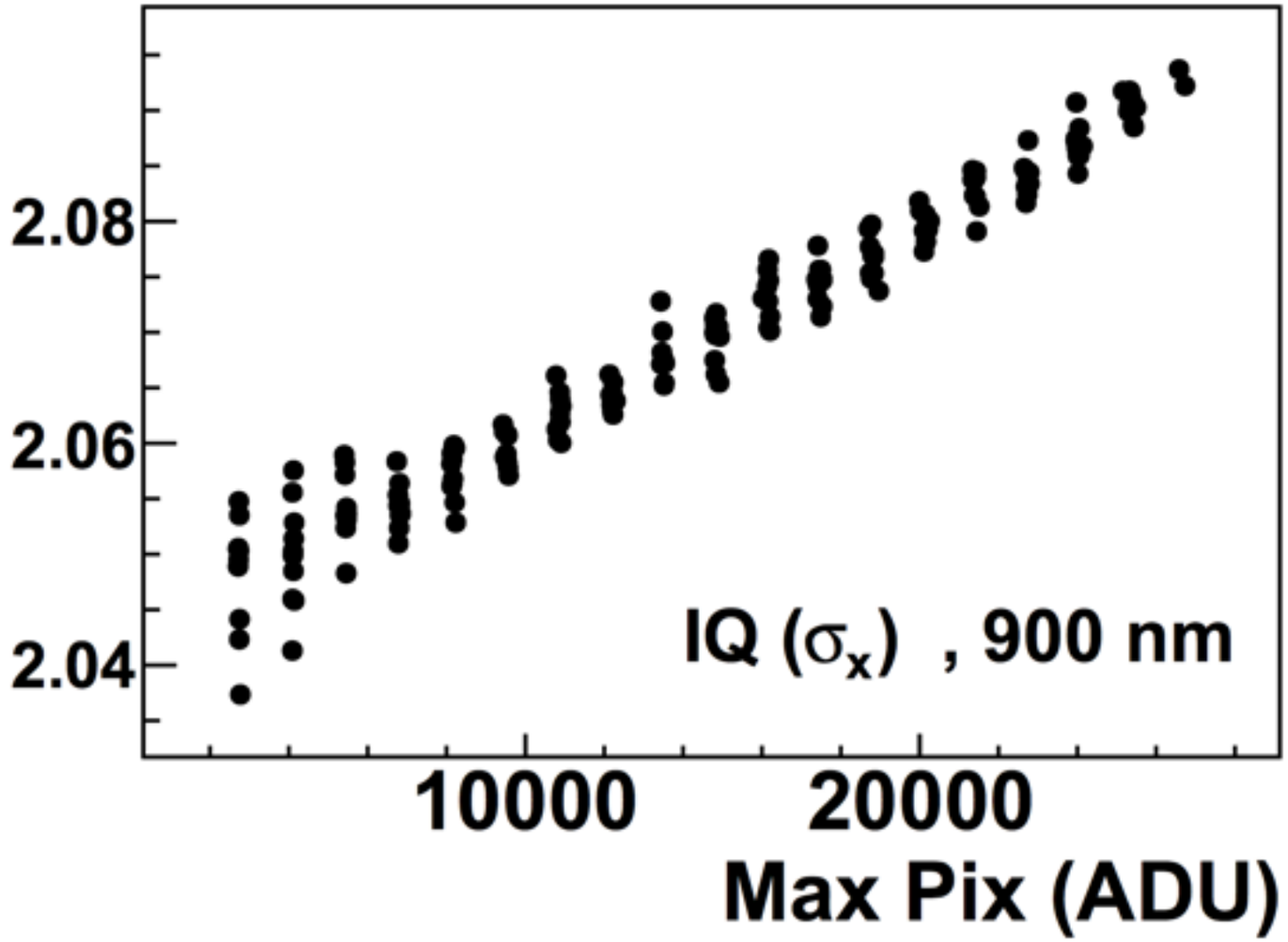


DARK ENERGY SURVEY

What is a Null Test?

- Any statistic that has a well-defined expected result can be considered (or recast as) a null test.

Image credit: Antilogus et al, 2014



Uses of Null Tests

- **Discovering Systematics**
  - If a null test fails, this may be a sign of some systematic error that isn't being corrected for in the analysis.
  - Therefore, it is a good idea to run a large suite of null tests, even if you don't have any a priori reason to think they might fail.
- **Testing Algorithms**
  - If you think your analysis pipeline is correcting for some systematic, a null test can help you determine if you are doing so correctly.
  - This can be done on simulated data as well as real data.
  - Extra null tests are possible on simulated data, since you can know the right answer, which is not available for real data

DARK ENERGY SURVEY

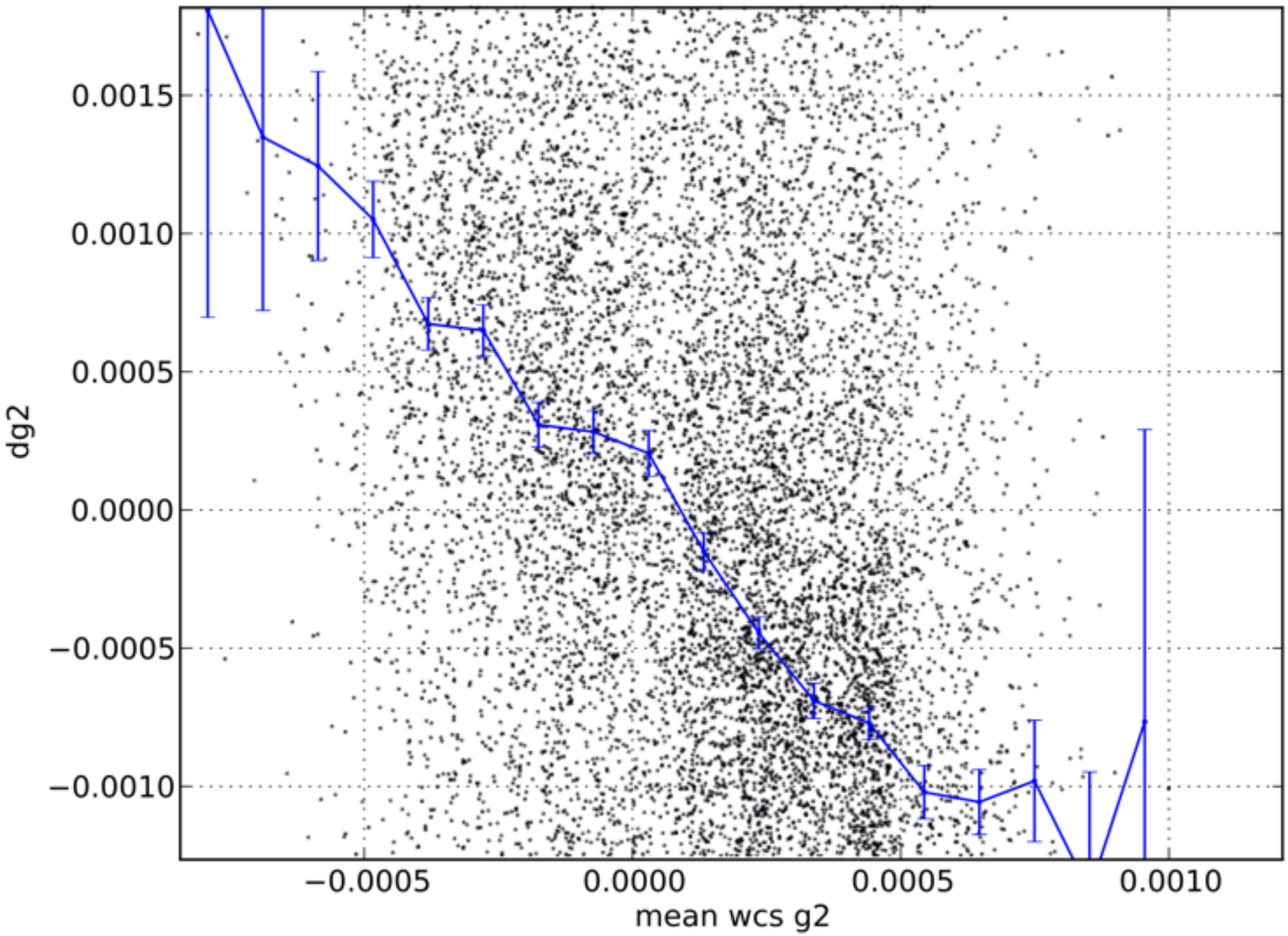
Null Test Philosophy

- If a null test is found to be non-zero, then there must be some uncorrected systematic error in the data.
- Null tests that are statistically significantly non-zero do not necessarily indicate a problem. Each null test needs a requirement based on the science to be done with the shears.
- Conversely, if all null tests pass, it does not imply that the data are free of systematics. It's hard to devise null tests that catch everything!
- Normally, a null test failure means some aspect of the analysis pipeline needs to be changed.

DARK ENERGY SURVEY

End-to-end Simulations

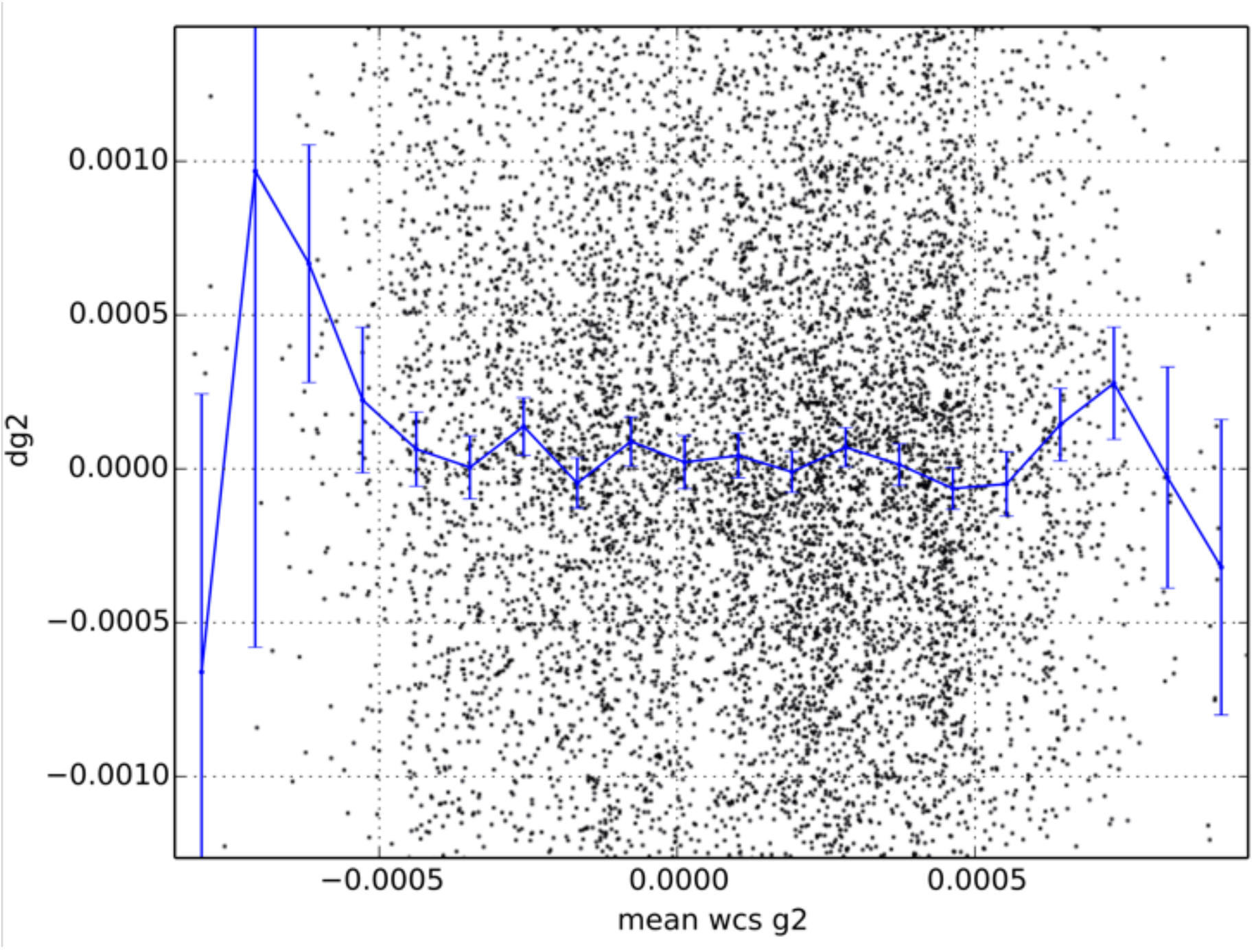
Bug in how the WCS was handled by im3shape



DARK ENERGY SURVEY

End-to-end Simulations

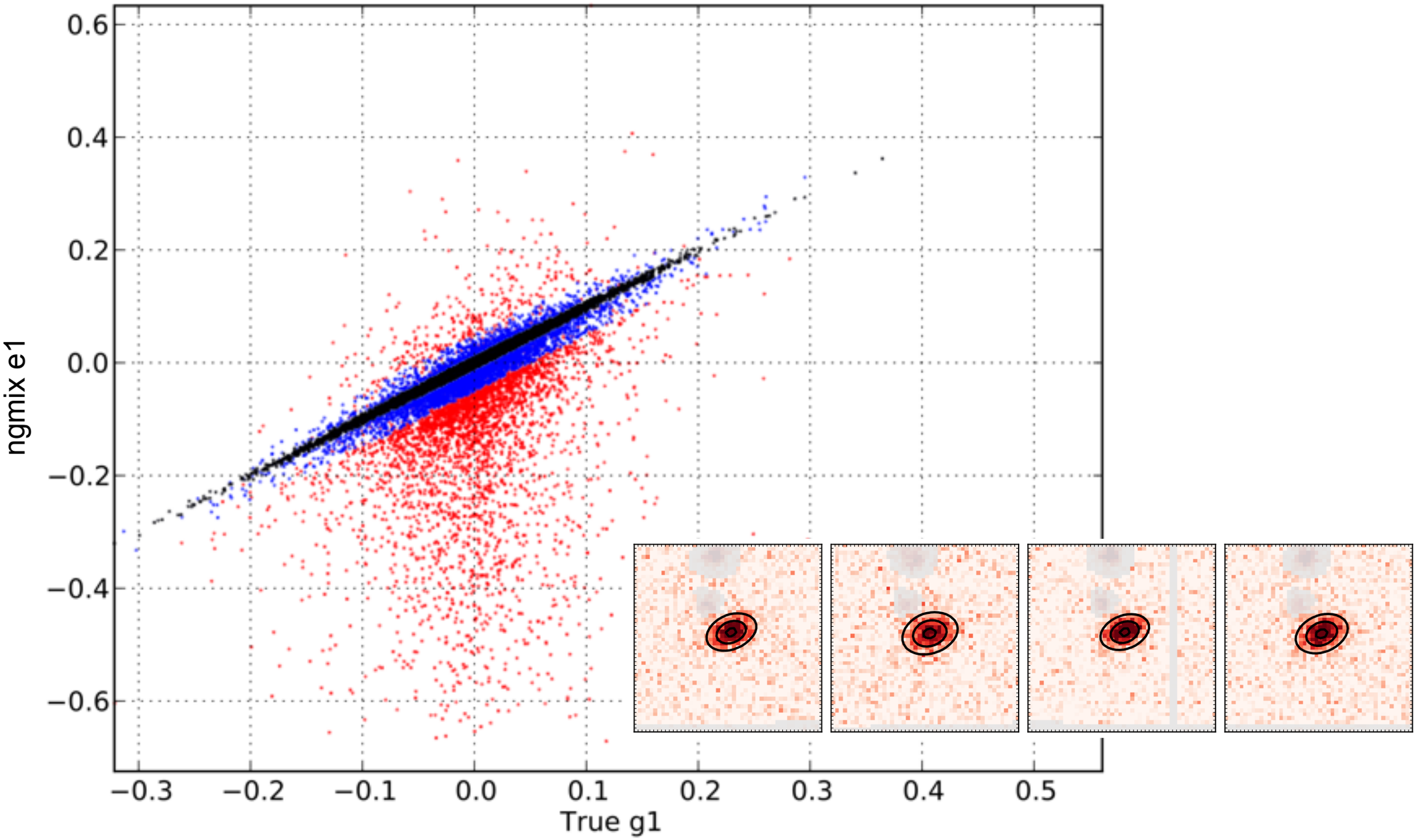
After bug fix



DARK ENERGY SURVEY

End-to-end Simulations

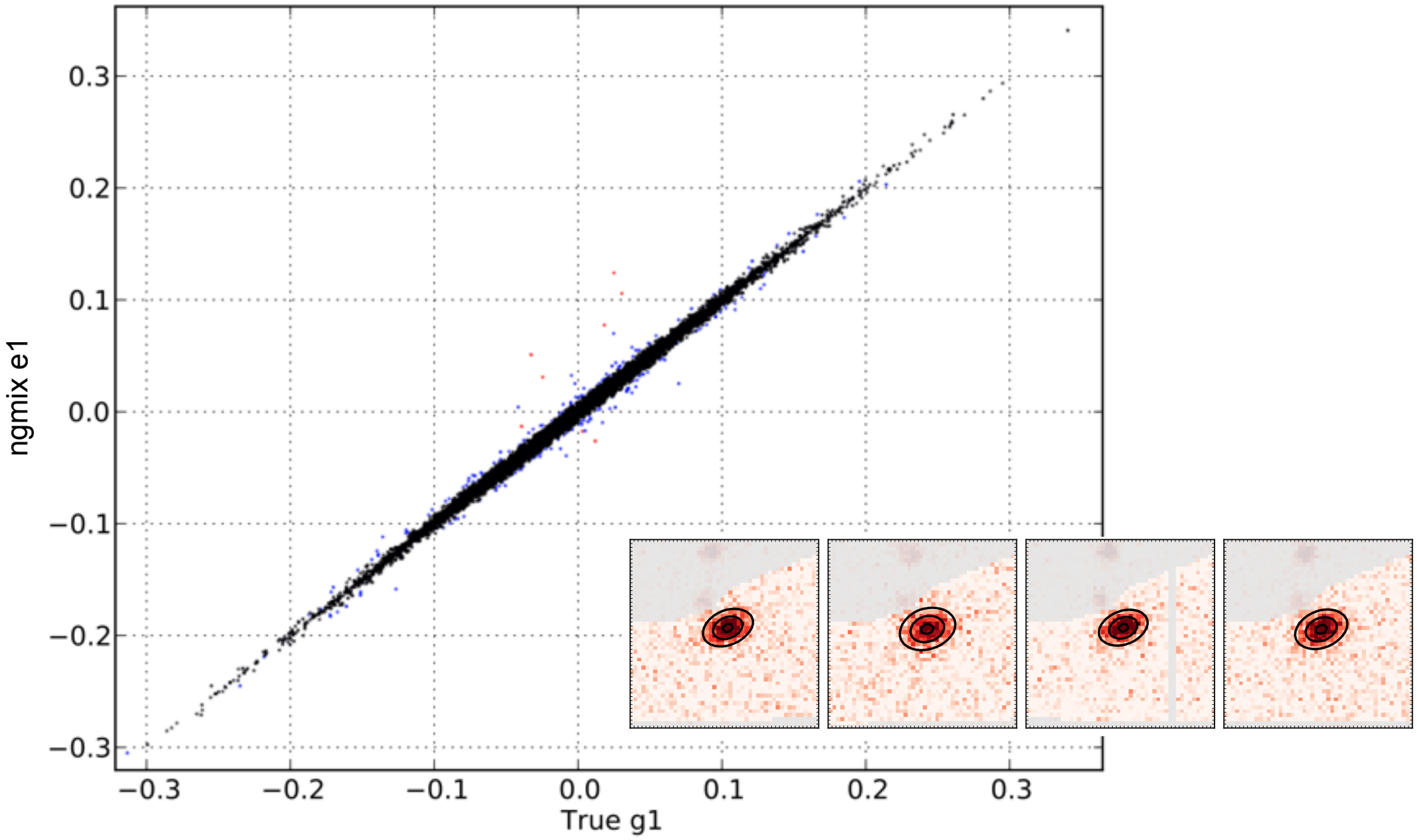
Mask using SExtractor segmentation maps



DARK ENERGY SURVEY

End-to-end Simulations

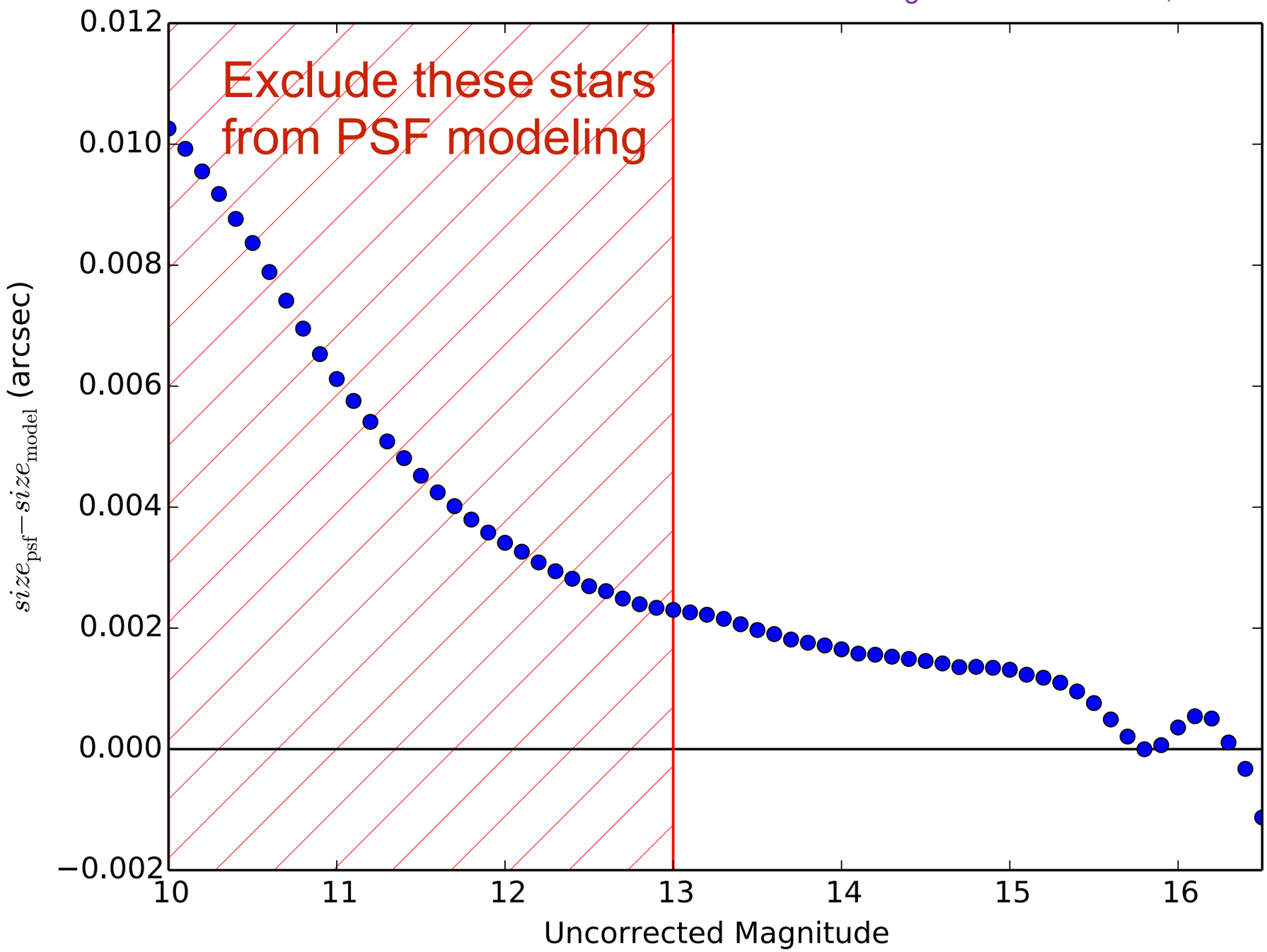
Mask using UberSeg



DARK ENERGY SURVEY

PSF Model Residuals

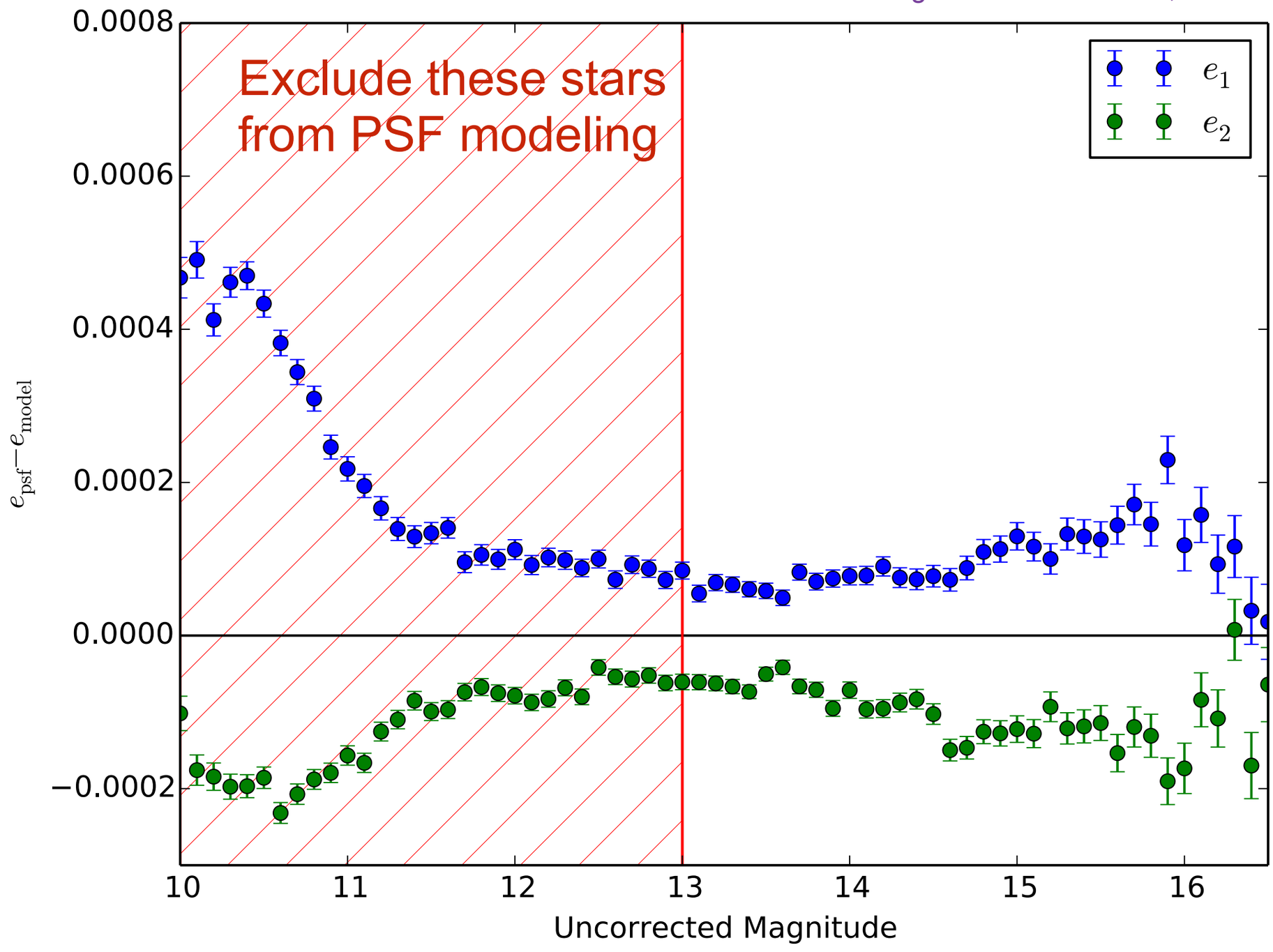
Image credit: Jarvis et al, 2015



DARK ENERGY SURVEY

PSF Model Residuals

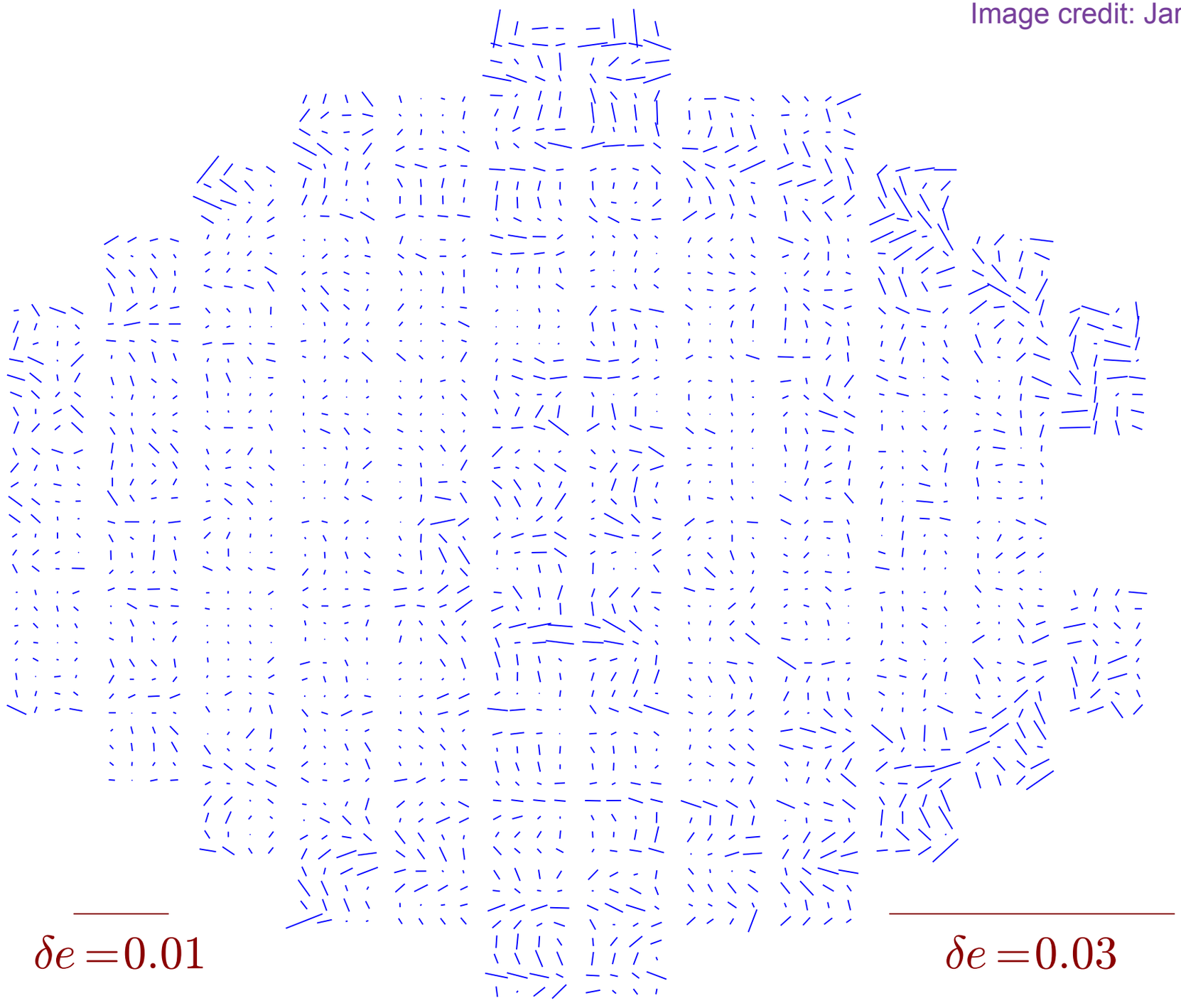
Image credit: Jarvis et al, 2015



PSF Model Residuals

DARK ENERGY SURVEY

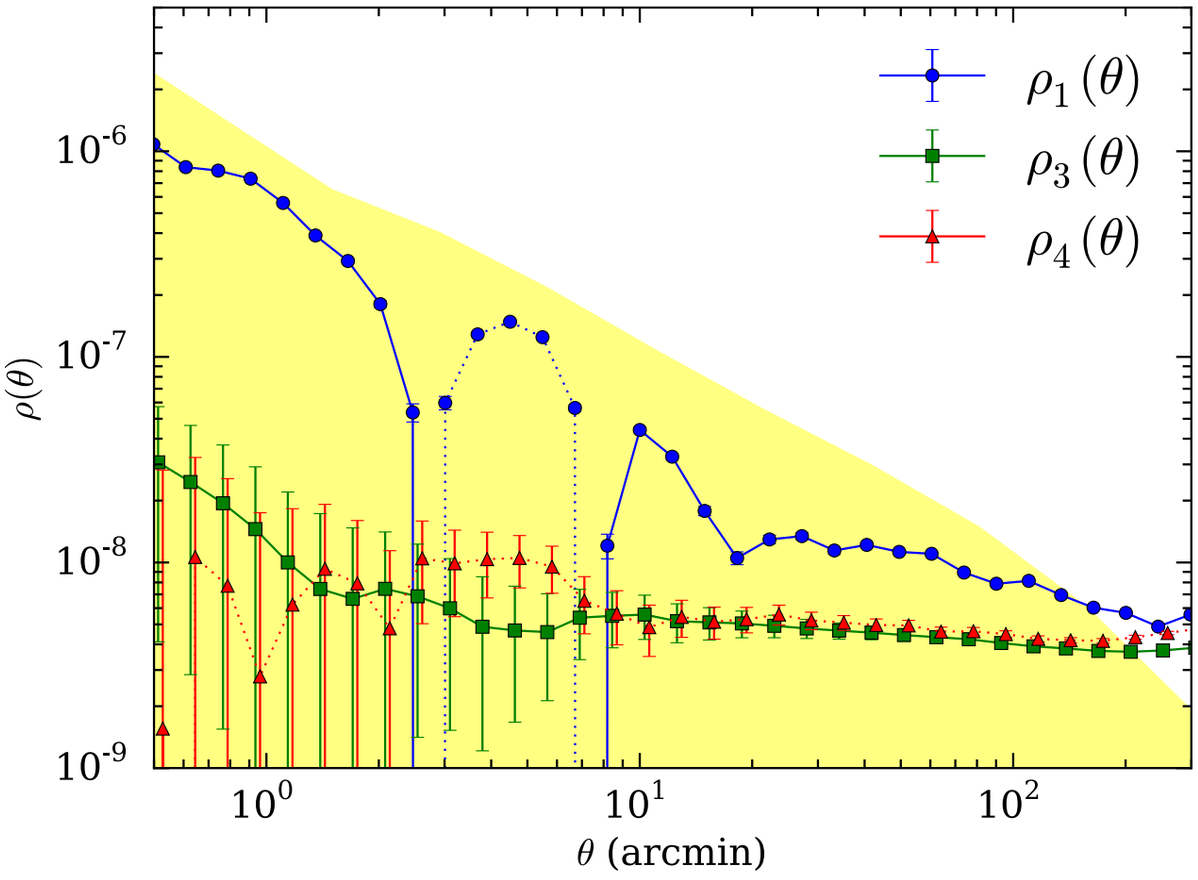
Image credit: Jarvis et al, 2015

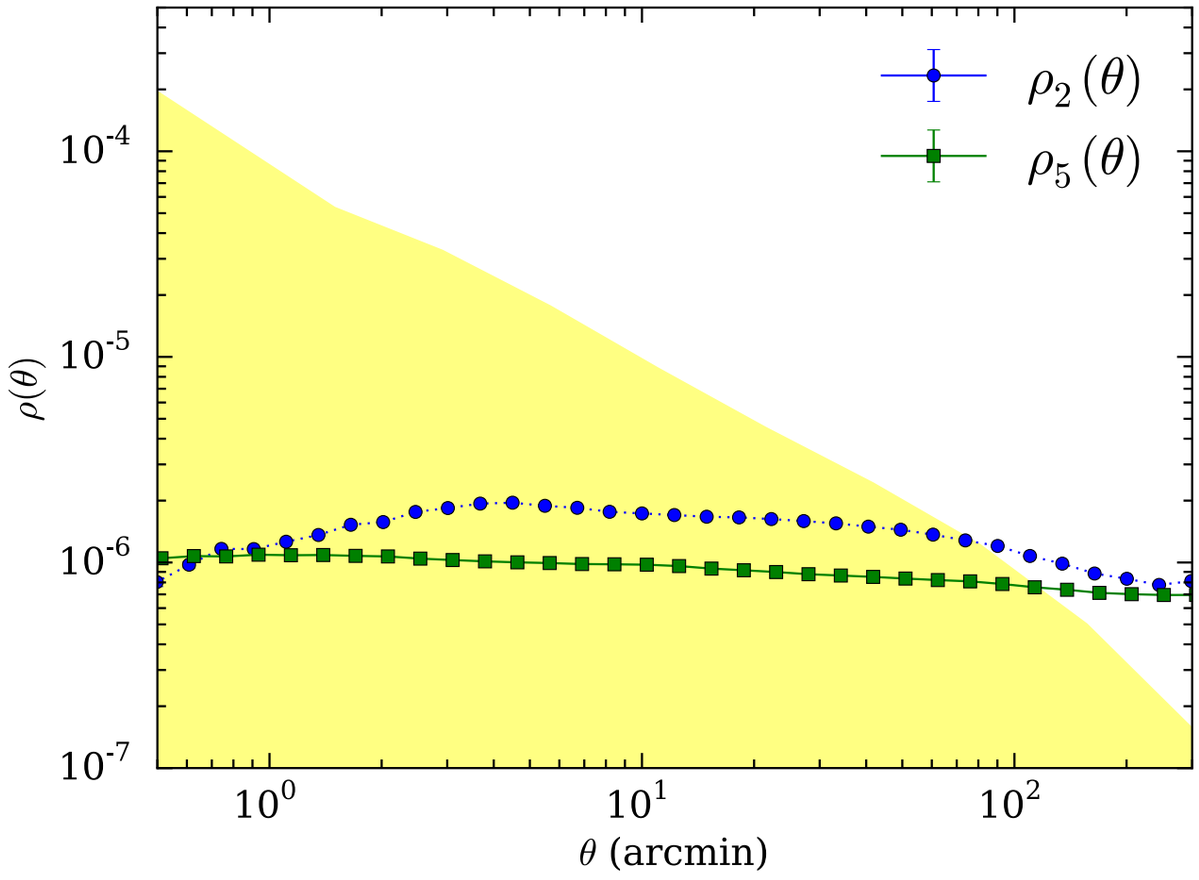


PSF Model Residuals

- ρ_1: Auto-correlation of PSF shape residuals
- ρ_2: Cross-correlation of PSF shape residuals with PSF shape
- ρ_3, ρ_4, ρ_5: Correlations involving PSF size residuals

Image credit: Jarvis et al, 2015

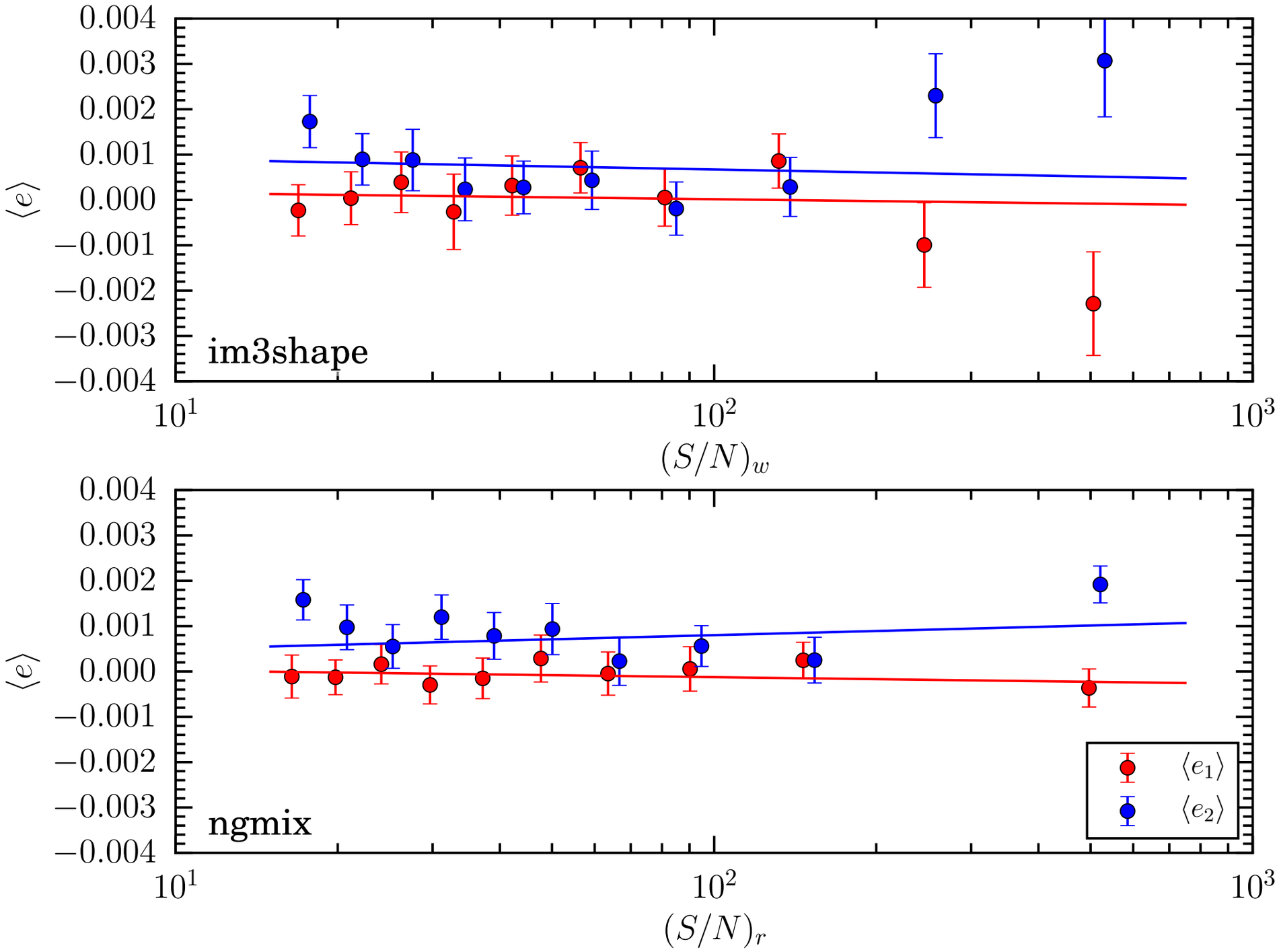




DARK ENERGY SURVEY

Shear vs Galaxy S/N

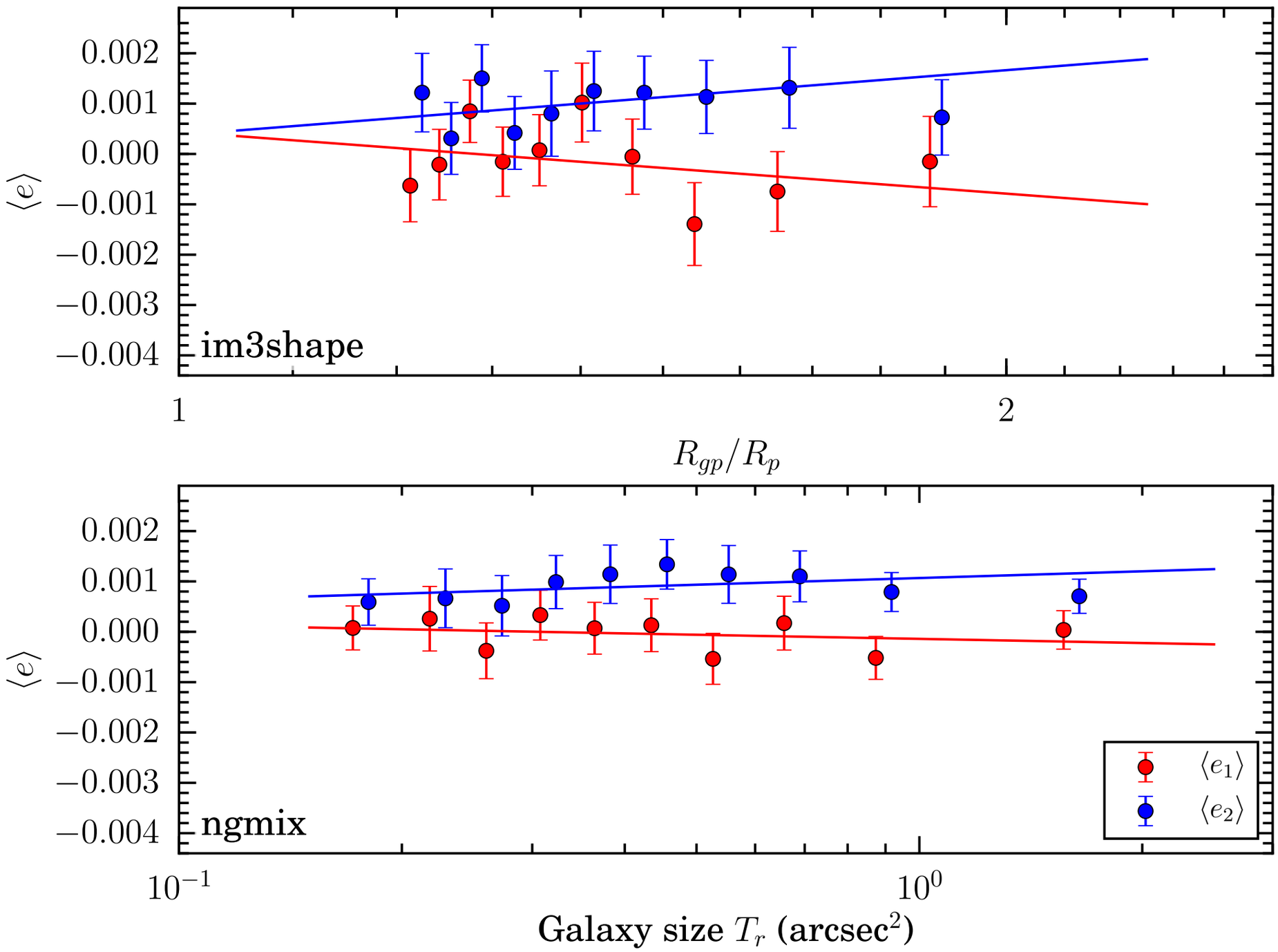
Image credit: Jarvis et al, 2015



Shear vs Galaxy Size

DARK ENERGY SURVEY

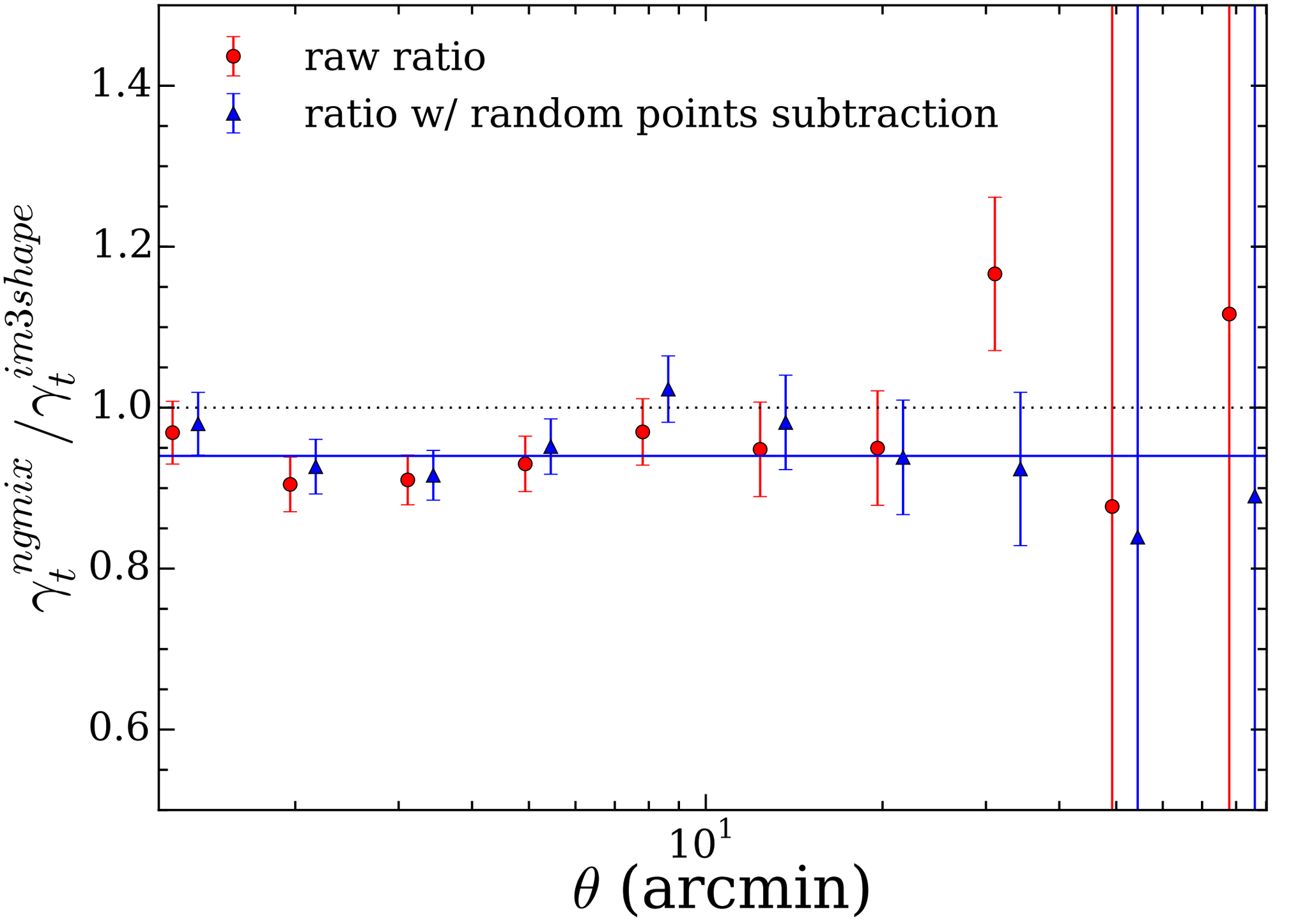
Image credit: Jarvis et al, 2015



DARK ENERGY SURVEY

Ratio of Signal Measured Using Two Algorithms

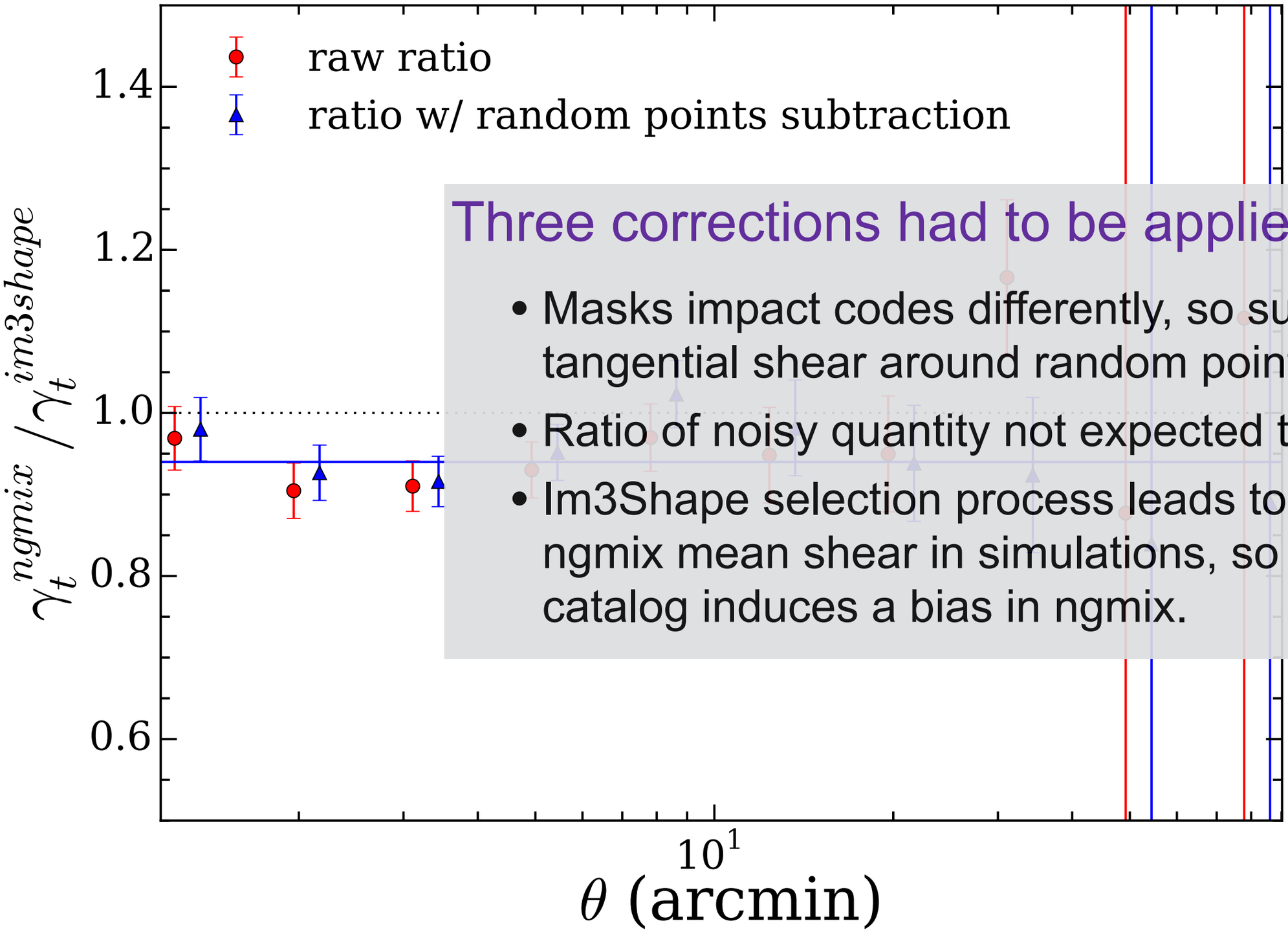
Image credit: Jarvis et al, 2015



Ratio of Signal Measured Using Two Algorithms

DARK ENERGY SURVEY

Image credit: Jarvis et al, 2015



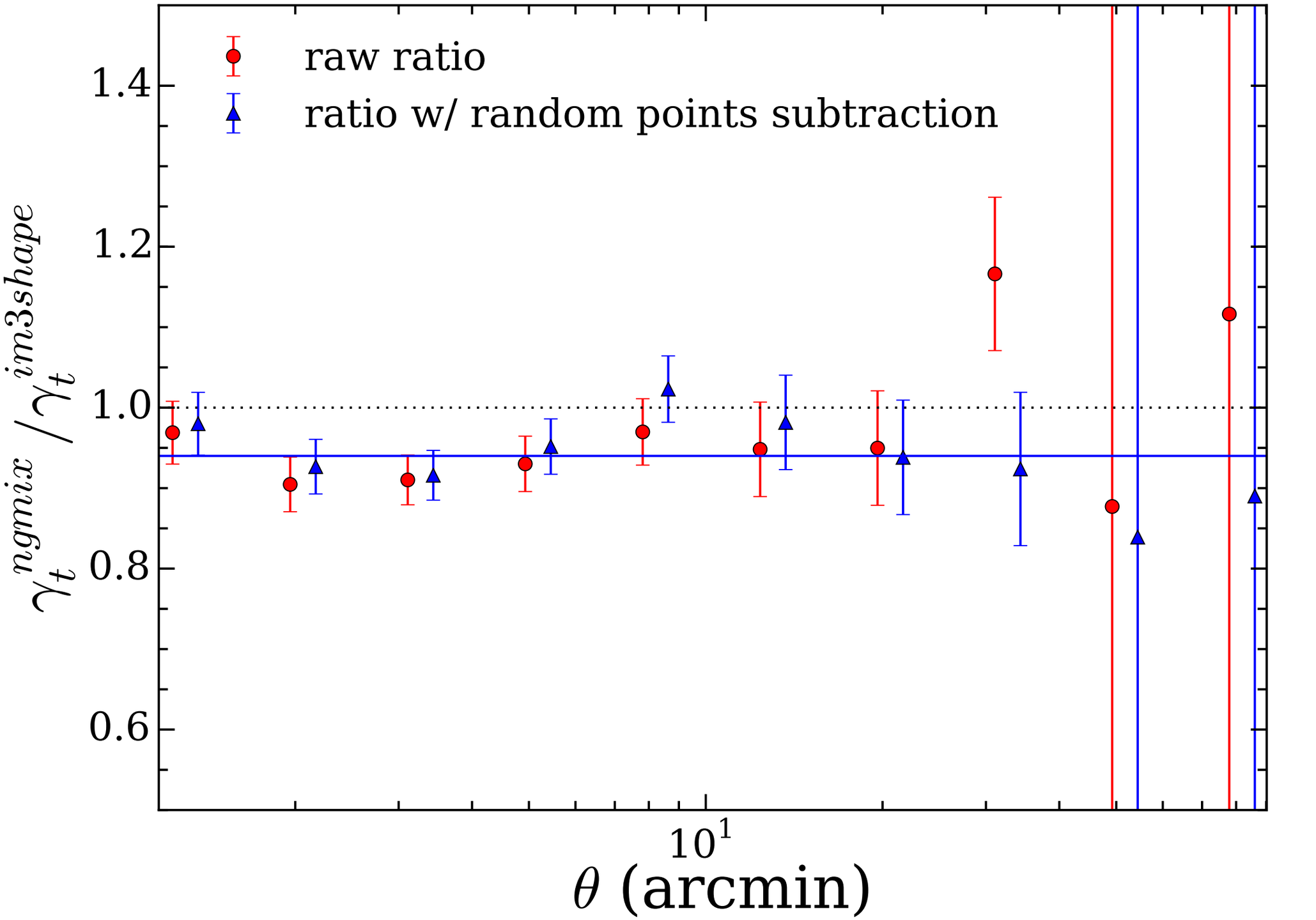
Three corrections had to be applied:

- Masks impact codes differently, so subtract tangential shear around random points.
- Ratio of noisy quantity not expected to be 1.
- Im3Shape selection process leads to bias in ngmix mean shear in simulations, so matched catalog induces a bias in ngmix.

DARK ENERGY SURVEY

Ratio of Signal Measured Using Two Algorithms

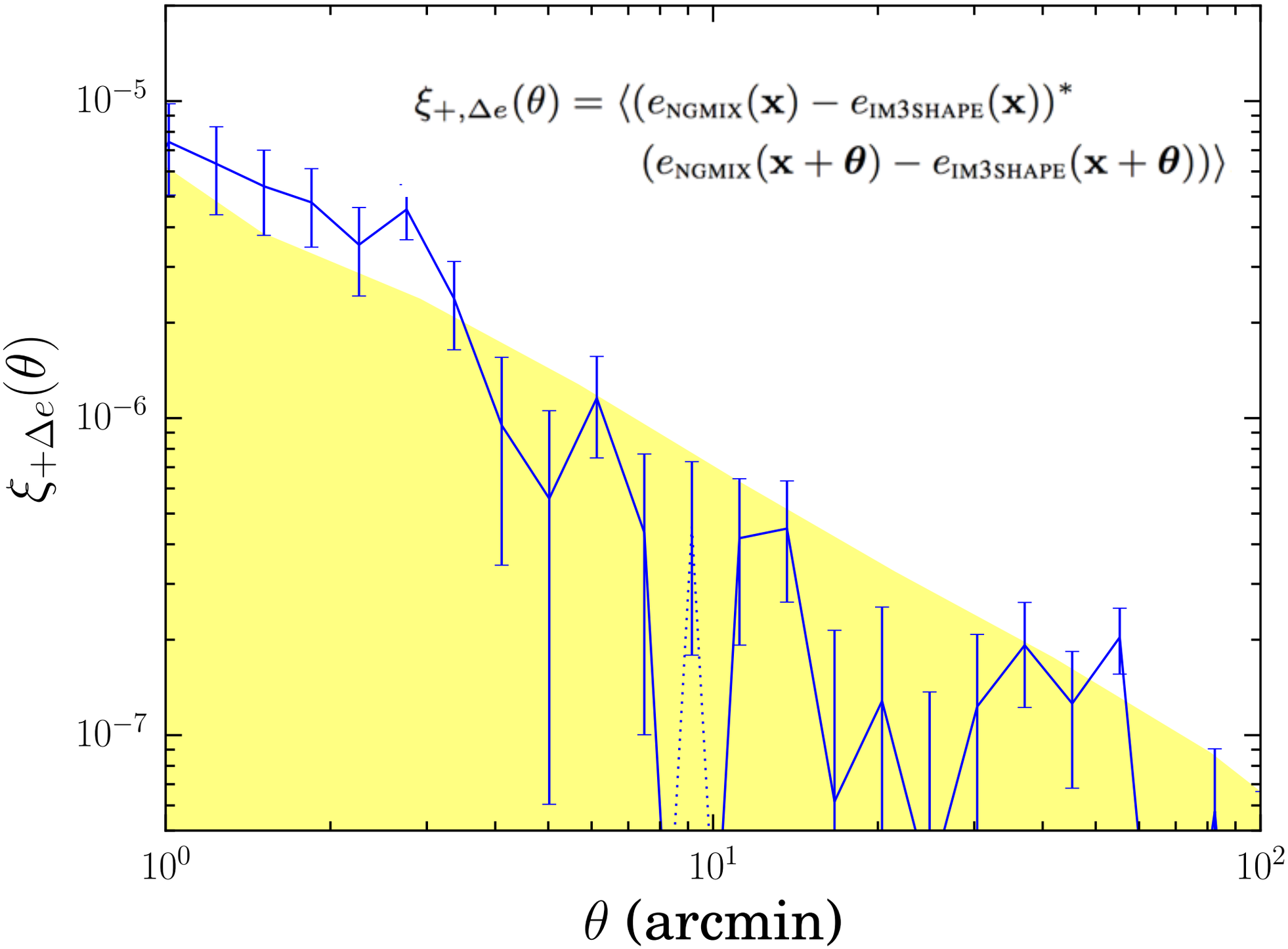
Image credit: Jarvis et al, 2015



Auto-correlation of Δe

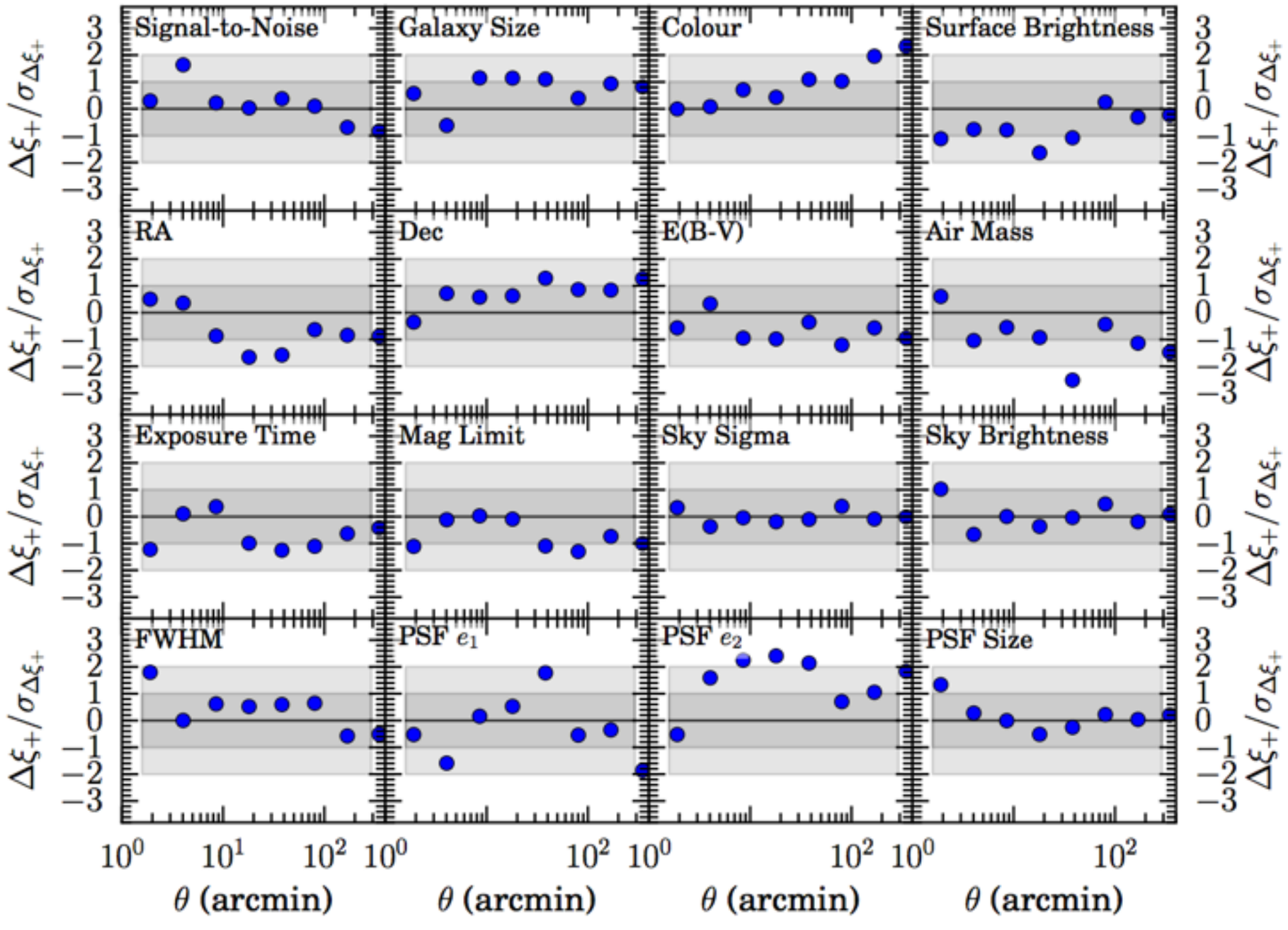
DARK ENERGY SURVEY

Image credit: Jarvis et al, 2015



Independence of Signal on Galaxy Properties

Image credit: Becker et al, 2015



Summary

- Routinely running a large suite of null tests can help discover problems in the analysis pipeline.
- For any null test where you bin the galaxies or split the galaxies in half, you need to be very careful about selection effects.
- Targeted null tests on simulations can be extremely valuable in verifying accuracy of analysis code.
- Most null tests can be run on commissioning data (ComCam) and other precursor data sets to get a head start on finding/fixing problems.